HUDSON, AGNES SCOTT. An Investigation of Skill Test Items as Measures of Golf Playing Ability. (1973) Directed by: Dr. Rosemary McGee. Pp. 71.

The purpose of this investigation was to devise golf skill tests that indicate golf playing ability in group testing situations. The skill tests were analyzed according to basic criteria for test selection.

The subjects were sixty students enrolled at the University of North Carolina at Greensboro. Forty-four students were enrolled in intermediate golf classes and sixteen students were volunteers with previous golf experience. A total of twenty-nine males and thirty-one females participated in the study. The tests were administered during the fall semester, 1972 and spring semester, 1973.

A drive test, an approach test, and a putting test were devised to indicate golfing ability. Distance and direction were the determining factors in assigning scoring values. Each test item consisted of twenty trials, scoring was done by partners, regulation balls were used, and the ball was scored where it came to rest. Different scoring values were assigned for men and women on the Five-Iron Drive Test.

To determine objectivity and administrative feasibility, an analysis was made according to stated criteria. To determine reliability, the Split-Halves method was utilized and coefficients obtained by the application of the Pearson Product-Moment method of correlation. The Spearman-Brown Prophecy formula was used to determine coefficients for entire tests and to predict coefficients

for additional trials. To determine validity, game score averages were correlated with test scores. Validity coefficients and intercorrelation coefficients were obtained by the application of the Pearson Product-Moment method of correlation. The Doolittle method for multiple correlations was used to determine validity coefficients for test combinations.

The findings in this study may be summarized as follows:

1. The single best indicator of golf playing ability was the Five-Iron Drive Test. This test item was objective, reliable, valid, and administratively feasible. The Split-Halves reliability coefficient was .93 and the coefficient using twenty trials was .96. The validity coefficient, using game score averages as the criterion, was .71.

2. The Approach Test was objective, reliable, and administratively feasible. The reliability coefficient using twenty trials was .88 and the validity coefficient was .66.

3. The intercorrelation coefficient for the combination of the Five-Iron Drive Test and Approach Test was .76 and the multiple correlation coefficient for the combination of the Five-Iron Drive Test and Approach Test was .73.

4. The reliability coefficient using twenty trials on the Putting Test was .65 and the validity coefficient was .59. The predicted number of trials necessary to reach an acceptable level of reliability (.81) was forty-five.

AN INVESTIGATION OF SKILL TEST ITEMS AS MEASURES OF GOLF PLAYING ABILITY

by

Agnes Scott Hudson

A Thesis Submitted to
the Faculty of the Graduate School at
The University of North Carolina at Greensboro
in Partial Fulfillment
of the Requirements for the Degree
Master of Science in Physical Education

Greensboro
1973

Approved by

Rosemary McGee
Thesis Adviser

APPROVAL PAGE

This thesis has been approved by the following committee of the Faculty of the Graduate School at the University of North Carolina at Greensboro.

Thesis Adviser Rosemary McGee

Committee Members Gail M. Tolonen

Norma D. Cocca

Elizabeth C. Umstead

November 14, 1973
Date of Acceptance by Committee

ACKNOWLEDGEMENTS

Special appreciation is extended to Dr. Rosemary McGee and Miss Ellen Griffin for their continuing influence on the school career of this student and for guidance pertaining to this study.

Appreciation is expressed to Miss Margaret Greene, Dr. Pauline Loeffler, Miss Nancy Porter, and Mr. James Swiggett for assistance in administrative processes and for the participation of students in their golf classes.

A sincere thank you is extended to those students who willingly volunteered their time and talents to participate in the study.

TABLE OF CONTENTS

CHAPTER

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Golf instructors frequently have access to a limited area for instruction and seldom are fortunate enough to have access to a golf course for class use. After the basics of the grip, stance, and swing are introduced, some measure of accuracy should be provided to enrich the learning experience. The most crudely constructed target area can aid the instructor and student in evaluating golfing ability since the elements of form and result can be related.

Often, locally prepared tests of skill ability serve as practice items and measurement items because standardized tests of skill ability are too time consuming to administer. This investigation was undertaken in an attempt to devise skill tests that would be statistically acceptable for evaluation purposes and suitable for group administration.

Statement of the Problem

The purpose of this study was to devise golf skill tests that indicate golf playing ability in group testing situations. A determination was made of the relationship of the skill test items as compared to acceptable standards of test selection in terms of objectivity, reliability, validity, and administrative feasibility.

Definition of Terms

The following definitions apply to this study:

1. Approach: A stroke with less than a full swing designed to cause the ball to travel a designated distance in the desired direction.

2. Drive: A stroke with a full swing designed to cause the ball to travel a maximum distance in the desired direction.

3. Putt: A light stroke on a putting green or simulated surface designed to cause the ball to travel a designated distance in the desired direction.

Assumptions Underlying the Research

1. A valid criterion of golf playing ability is actual game scores.

2. The fundamental elements of the game of golf may be classified as the drive, approach, and putt.

3. The characteristics of the drive, approach, and putt are distinguishable and a valid evaluation of golf playing ability should include measurements of each element.

4. A test of the drive requires an adjustment in distance and/or scoring in order to be appropriate for both men and women.

Scope of the Study

This study is limited in the following ways:

1. Outdoor skill tests of the drive and approach and an indoor skill test of the putt were devised by the investigator.

2. The subjects were male and female students who were enrolled at the University of North Carolina at Greensboro and had previous golf experience.

3. Game scores were obtained from play on the University golf course.

4. Test data were collected during the fall semester, 1972 and spring semester, 1973.

Significance of the Study

Golf skill tests are a necessity when evaluations are desired that are not subjectively based and when testing on a golf course is not practical. Since many skill tests that are statistically acceptable for evaluation purposes are not suitable for use in group testing situations, several factors of administrative feasibility were considered in this study. Very often, skill tests that are designed to measure golfing ability include only one test item. This study was conducted on the assumption that a valid evaluation of golfing ability should include measurements of the drive, approach, and putt. A review of the literature did not reveal a three-item battery of the drive, approach, and putt with acceptable standards of objectivity, reliability, validity, and administrative feasibility.

The criterion used for validation purposes, in those reported tests establishing a validity coefficient, was usually judges' ratings. In an attempt to more accurately determine validity, actual game scores of the subjects were used in this study.

CHAPTER II

REVIEW OF LITERATURE

Skill tests for golf first appeared in the literature in the 1930s. Although an abundance of the "how to" type books relating to techniques and strategy are available, developing scientific measures to assess golfing ability has been confined almost completely to research undertakings in the school environment. The review of literature in this study consists of (1) ascertaining the criteria for test selection, (2) determining the place of skill testing in sport, and (3) reviewing previous tests that have been constructed to measure golf playing ability. Related literature was reviewed that reported the use of some form of golf testing in conjunction with the main part of a study. For the most part, these tests were developed locally to serve a specific purpose and a determination was not made about possibilities for further use. The main purposes in reviewing golf skill tests were to determine the relationship of the test items as compared to acceptable standards of test selection and to consider procedures that might be useful in constructing or revising skill tests.

Criteria for Test Selection

Authors of measurement texts are pretty much in agreement about the factors that should be considered when selecting tests. References are most frequently made concerning the factors of statistical evidence

and administrative feasibility. Scientific authenticity refers to the criteria used to evaluate in terms of scientific worth and includes measures of validity, objectivity, reliability, and norms (Mathews, 1968). According to Barrow and McGee (1971), validity refers to the honesty of a test--the degree that a test measures what it describes as measuring; reliability refers to the degree with which performance is repeated when the test is administered under similar conditions; objectivity is a measure of the agreement with which various persons score the same test; and norms are scales reflecting standards of achievement that allow for comparison of performance among members of a population. Norms become a consideration in test construction when acceptable standards for the other technical criteria have been met.

Factors to be considered for administrative feasibility relate to requirements in terms of time and money. Neilson and Jensen (1972) stated that only those valid measures that can be administered to a group of students at one time and require little equipment are highly useful in physical education.

In addition to the above considerations, some authors suggest other criteria that are necessary in test selection. Scott and French (1959), for example, suggested that tests should measure important abilities, be game-like, encourage good form, involve one performer, be interesting and meaningful, be of suitable difficulty, differentiate between levels of ability, provide for accurate scoring, and allow for a sufficient number of trials to eliminate chance deviations.

The Place of Skill Testing in Sport

Test results are useful as devices to aid in grading, classification, guidance, motivation, and research (Barrow and McGee, 1971). The use of skill test results for predictive purposes, however, is limited in several respects. Lawther (1968) stated that skill tests with a high enough validity to be used in individual selection or for prediction are rare because of three major problems: (1) rate and amount of improvement can seldom be determined because of the low relationship between initial and final scores; (2) certain factors in performance such as personality, attitude, and adjustments to stress are impossible to measure; and (3) specific test items in isolation do not represent the same performance as when integrated into the total activity.

Pertaining to golf skill testing specifically, Willgoose (1961) explained the lack of standardized tests since the game score card is an appraisal instrument in itself. He commented on the ease of constructing local tests of the putt, drive, and chip to a green if such testing is desired. Scott and French (1959) recommended the use of the game score card as a test score, with nine holes of play being desirable. They recommended that game scores be supplemented by ratings since the playing of games is so time consuming and the performance of beginners is usually erratic.

The relative importance of the different elements that comprise the game of golf is a topic that can be debated indefinitely. Some objectivity, however, is present in the reported tabulation of strokes by a foursome of experts who found the following percentages:

drive--30 percent, approach--30 percent, and putt--40 percent (Coffey, 1946). Perhaps the numerical importance of putting is put into its proper perspective with the observation by Lema (1966:90): "Considering how irrelevant the putting stroke is in any earnest study of golfing technique, it's ironical that this narrow part of the game should account for about one-half your strokes!"

Golf Skill Tests

The main purposes in reviewing reported skill tests were to determine the relationship of the test items as compared to acceptable standards of test selection and to consider procedures that might be useful in constructing or revising skill tests. Remarks that summarize certain tests or groups of tests are in reference to basic criteria for test selection. Tests are reported in the order of their appearance in the literature within the respective headings of (1) putting tests, (2) approach tests, (3) drive tests, and (4) test batteries.

Putting Tests

Lumpkin (1945) devised a putting test in order to estimate the skill level of the golfer. The test consisted of putting sixteen balls each from distances of five, ten, twenty, and thirty feet for a total of sixty-four trials. The target area was a cup on a level outdoor green and the score for each trial was the total number of putts required to hit the ball into the cup. The reliability coefficient for forty-two beginners was .74. Although considered

satisfactory for group measurement, the test was not found to be administratively feasible since only six or seven subjects could complete one-half the test in one class period, with several subjects being tested simultaneously.

Kelliher (1963) constructed two tests in order to compare the effectiveness of two styles of putting. In test A, the alignment ability test, five putts each were taken from distances of six, fifteen, and thirty feet from the hole. Putts were scored "in" if the ball crossed any part of the hole, regardless of distance. If the ball missed the hole, distances to the right and left were recorded in inches. In test B, the alignment ability plus distance judgment test, putts were measured in inches from the ball's stopping point to the hole. Reliability and validity were not reported.

Gibb (1964), in conjunction with a study to determine the value of the concentrator putter as a practice device, utilized three putting tests similar to those used by Kelliher. Each test consisted of twelve putts at distances of five, fifteen, and thirty feet. The best and poorest putts were eliminated at each distance. Subjects rotated so putts were not taken from the same distance in succession. In the distance ability test, a rectangle four feet long and four and one-half inches wide was the target area. If the ball did not stop in the rectangle, the distance deviation was recorded in inches. In the alignment ability test, a three-inch area to the right and left of the cup was indicated. The ball was scored if it passed any part of the hole. The putt was recorded as a near miss if the ball rolled in the three-inch area to the right or left of the

hole. If the ball missed the cup, the distance to the right or left was recorded in inches. In the distance plus alignment ability test, putts that did not go into the hole were measured by the distance in inches from the hole. Reliability and validity were not determined.

Neale and Anderson (1966) utilized a device that measured accuracy of aim in order to compare conventional and croquet-style putters. The subject's ability to line up a shot in the designated direction was determined by a device to which the putter was attached and another device that determined the angle of the clubface in preparation for a putt. This test would offer limited information on putting ability as it measured accuracy of aim only.

Bowen (1968) designed a test to compare putting errors of golfers using different points of aim. Twenty-five shots each were taken from distances of fifteen, twenty-five, and thirty-five feet on a carpet designed for an outdoor area. Results were recorded for distance errors to the nearest inch and directional errors by left, right, long, and short quadrants. Lines one inch wide were drawn on each side of the cup extending to six feet in order to determine directional errors. Statistical tables were not included with the report of this study.

Randleman (1969) constructed a putting test in order to study the effectiveness of two styles of putting. He utilized a regulation level green with ten trials each taken from distances of five, fifteen, and twenty-five feet. Measurements were taken, after each putt, of distance and accuracy to the nearest inch. An alignment score was then calculated from the distance and accuracy scores. Low

reliabilities were reported and the investigator recommended using five, twenty-five, and fifty feet as distances in a future study. He also recommended changing the putting distance after each trial, counting putts that rimmed the cup as putts made, and using a simulated surface inside.

Williford (1970) designed an indoor putting test and established the reliability and validity of the indoor test and an outdoor putting test used by golf instructors at the University of North Carolina at Greensboro. The indoor test was performed on an artificial green constructed by the investigator. An indoor-outdoor carpet placed on a styrofoam sub-surface was cut so that five holes of varying distances, breaks, and elevations served as the target area. The score for the entire test consisted of the total number of strokes necessary to putt the five holes. The criteria for validation purposes were game play putts and an outdoor putting test. The outdoor putting test was conducted on a practice green with nine holes. Three holes were six feet in distance, three were twelve feet, and three were eighteen feet. The score for the entire test was the total number of strokes necessary to putt the nine holes. The investigator found that each of the reliability and validity coefficients for 171 beginners on the indoor and outdoor tests represented unacceptable correlations.

Approach Tests

Watts (1942) devised targets for testing the approach shot at distances of ten to fifty yards. Based on the principles of exactness

of force and direction, the target areas were designed in the shape of trapezoids. The investigator found that the limited number of ten subjects did not allow for conclusions to be established.

Nelson (1967) constructed a pitching test to measure golf achievement. The subjects selected a club and took ten trials, after three practice hits, at a target consisting of seven concentric circles; the outer circle had a diameter of sixty-six feet. A restraining line was placed twenty yards from the center of the target and a hitting line was placed forty yards from the center. In order to score, balls has to be airborne until past the restraining line. Scores ranged from two to ten points per trial. A validity coefficient of .86 was obtained when related to judges' ratings and .79 when correlated with golf scores. A reliability coefficient of .83 was obtained by the test-retest method. Identifying information about the subjects participating in the testing or the conditions under which the coefficients were obtained were not specified in the report of this study.

Smith (1968) utilized a regulation outdoor green for a pitch-and-run test to be used in conjunction with a study of the effectiveness of television video tape instant playback. The seven iron was used and fifteen trials were taken at a distance of seventy-five feet from the pin. The target area consisted of fifteen concentric circles; the outer circle had a diameter of thirty feet. Scores ranged from one to fifteen points per trial with sixteen points for a hole-in-one. The reliability coefficient for thirty-seven college women classified as beginners was .83. Low validity was found between judges' ratings and test scores.

West and Thorpe (1968) constructed an eight-iron approach test in which two practice shots and twelve trials were taken on separate days for a total of twenty-four trials. The target area consisted of six concentric circles on a level, grassy area; the outer circle had a diameter of eighteen yards. Subjects stood behind a restraining line placed twelve yards from the pin. Each trial was scored by two methods, consequently, a total of twenty-one points was possible on each trial as a result of a flight times accuracy score. The vertical angle of projection (flight) was rated by the test administrator as follows: (1) topped ball--one point, (2) low angle of projection--twenty-nine degrees or lower--two points, or (3) high angle of projection--thirty degrees or higher--three points. The accuracy score was measured according to where the ball came to rest in the circles and ranged from one to seven points per trial. The reliability coefficient for 424 college women classified as beginners was .75 and logical validity was claimed.

Drive Tests

McKee (1950) developed a test using the five iron and the two iron to measure the full swing. The test consisted of twenty trials into an area lined with a rope 175 yards long and marked with colored ribbons at 25-yard intervals and white ribbons at 5-yard intervals. Stakes were used to mark the spot where the ball first hit. A ball in the air less than six-tenths of a second did not count as a trial. Measurements were taken for distance along the intended line of flight

by the use of the rope and the deviation from the intended line of flight was measured by the use of a steel tape placed at right angles to the rope. From these two measurements, the range, velocity, angle of impact, and angle of deviation were calculated. For the two iron, reliability coefficients for thirty women subjects were .92 for the range, .86 for velocity, .81 for angle of impact, and .82 for angle of deviation. For the five iron, reliability coefficients for the same subjects were .95 for the range, .89 for velocity, .89 for angle of impact, and .60 for angle of deviation. Validity was assumed to be inherent in the test.

Reese (1960) reported an outdoor and an indoor test to measure the ability to hit with the five iron. The outdoor test consisted of three practice hits and twenty trials in succession. Target areas were marked by flags placed twenty, forty, and sixty yards from the hitting line. Trials were scored according to where the ball first hit; values were seven points if beyond the sixty-yard line, five points if between the forty- and sixty-yard lines, and three points if between the twenty- and forty-yard lines. One point was scored if the ball was hit beyond the twenty-yard line but not as high as the subject's head. The reliability coefficient for 109 college women beginners and 16 intermediates was .89 and self validity was claimed. The indoor test consisted of hitting plastic balls into a target area adapted to local conditions. The reliability coefficient for twenty trials was .92. The outdoor test correlated .50 with the indoor test which led the investigator to conclude that the outdoor and indoor tests did not measure identical abilities and that one test could not be substituted for the other.

Benson (1963) developed a test of golf ability using the five iron. Five practice shots and twenty trials were taken at a target area designated by distance signs placed from 25 yards to 150 yards at 25-yard intervals. Parallel to the hitting line and 150 yards away were nine areas marked off in 5-yard increments on each side of the center. The scorer stood behind the hitter and recorded (1) an estimate of the distance the ball traveled in flight and (2) the deviation from a straight line extending to the 150-yard marker. The average of twenty trials provided a distance score and a deviation score. The reliability coefficients were .90 for distance and .70 for deviation. The distance and deviation scores in combination produced a multiple correlation coefficient of .94 using golf scores as a criterion of validity. Identifying information about the subjects was not specified in the report of this study.

Bevacqua (1964) modified the outdoor test reported by Reese in a study on club progressions. Scoring zones were extended to include an eighty-yard line and a one hundred-yard line. A ball that hit between the eighty- and one hundred-yard lines was scored nine points and a ball that hit beyond the one hundred-yard line was scored eleven points. Bevacqua eliminated the requirement for the scorer to decide whether the ball went as high as the subject's head. The reliability coefficients for 105 beginners were .80 with the five iron, .76 with the eight iron, and .85 with the three wood. Three practice shots and twenty trials were taken on each application of the test.

Test Batteries

Clevett (1931) experimented with indoor tests in some of the initial testing reported in golf. His brassie and midiron tests each consisted of hitting ten shots at a vertical target twenty-one feet from the hitter. The target was hung at the end of a net and divided into twenty-five areas, each of which was twenty inches square. The mashie test consisted of hitting ten shots at a target of mats on the floor with the hitter standing fifteen yards from the nearest edge of the target. The target was divided into twenty-five areas, each of which was four feet square. The putting test consisted of hitting ten shots on a carpet twenty-seven inches wide and twenty feet long. The carpet was divided into forty-eight areas, each of which was nine inches square. In the putting test, as in the other tests, scores ranged from one to ten points per trial. A circle the size of a regulation hole was painted on the carpet and the hitting line was fifteen feet from the cup. Clevett observed that the tests were easily administered, interesting to beginners and more advanced players, accurate, useful for teaching, and distinctly objective. He believed that further use would demonstrate the validity and reliability of the tests, although reliability and validity were not determined from the data collected by Clevett.

Wood (1933) investigated the use of a driving cage for a golf target and test. The investigator found that different targets should be set up for the brassis and mid-iron clubs based on the areas of greatest concentration of shots.

Autrey (1937) constructed outdoor and indoor tests to measure golf playing ability. Only the outdoor distance driving test was found to be reliable. It consisted of taking ten trials with a two wood on an outdoor driving range labeled up to 200 yards. The distance to the nearest five yards was determined from where the ball stopped rolling. The reliability coefficient for forty-two subjects using twenty trials was .84.

Kelly (1944) developed a battery of tests to measure the golfing ability of beginners. Four approach tests and one distance test were administered.

The twenty-five-yard approach test consisted of hitting ten shots with either a nine iron or seven iron at a target two-by-three yards on the twenty-five-yard line. Each trial was scored by the deviation in yards from the target, both laterally and vertically. For all tests, the players were hitting into a lined field of lanes three yards wide.

The fifty-yard approach test was similar to the twenty-five-yard approach test except that a seven iron or five iron was used. The seventy-five-yard approach test allowed for the use of a five iron or four iron. For the one hundred-yard approach test, a five iron or two iron was used. For the two-iron distance test, the ball was hit as far as possible and the total distance was recorded with yardage subtracted for deviation to the right or left of the subject's three-yard hitting lane.

All the tests were determined to be reliable as scored. The twenty-five-yard approach test was the most reliable. The coefficient

of .92 was computed using the best two even trials and the best two odd trials. A study of validity used a six-hole playing test as the criterion. The two-iron distance test correlated best with the criterion with a coefficient of .70. The author recommended further study to establish validity since the six-hole playing test offered a limited number of holes and putting was omitted in scoring due to the poor condition of the greens. Ninety college women in beginning or intermediate golf classes participated in the study.

Coffey (1946) constructed a battery of indoor tests using regulation hard balls. Those tests with the highest reliabilities are described in this study. The Drive I test utilized the spoon club and a canvas wall target placed twenty feet from the center of a cocoa mat. The target area consisted of five concentric circles with areas numbered from one to six. The outer circle was ten feet in diameter. The approach test utilized the mashie club and a floor target with its center twenty yards from the center of a cocoa mat. The target area was twenty feet by forty feet and divided into ten scoring areas numbered from one to five. The ball was scored according to where it first hit the target. The Putt I test utilized a level carpet and a putting cup. The target area consisted of five concentric circles; the outer circle had a diameter of thirty-two inches. Areas were numbered from one to six and the cup was nine feet from the hitting line. Reliability was computed using thirty trials. The reliability coefficient for the Drive I test was .712; the approach test, .787; and the Putt I test, .509. Validity, using judges' ratings as the criterion, was

the highest for the approach test with a coefficient of .670. Sixty-seven high-school girls participated in the study.

Vanderhoof (1956) attempted to measure golfing ability by a drive test, a five-iron test, and a seven-iron test. Testing was done in an indoor area and plastic balls were used. The drive test and five-iron test were similar except that the five-iron test did not utilize a tee. Fifteen trials were taken with the two wood or five iron at a ten-pin target placed behind the testing area. In order to score, the ball must have passed under a rope eight feet high and fourteen feet from the hitter and land in the testing area. One point was scored for the area fourteen feet from the hitter and twenty feet in length. Two points were scored for the area thirty-four feet from the hitter and twenty feet in length. Three points were scored for the area fifty-four feet from the hitter and twenty feet in length. For the seven-iron test, the ball must have passed over a rope which was two and one-half feet high and fifteen feet from the hitter. The testing area consisted of two circles; diameters were six feet and twelve feet. The center of the target area was twenty-five feet from the hitter.

The reliability coefficients for 110 college women were .90 for the drive test, .84 for the five-iron test, and .86 for the seven-iron test. A subjective rating of judges proved to be more satisfactory than scores from playing six holes of golf as a criterion for validation purposes. Validity coefficients, using the judges' ratings, were .71 for the drive test, .66 for the five-iron test, and .58 for the seven-iron test. The combination of the drive test and five-iron test produced a multiple correlation coefficient of .78.

Olsen (1958) developed tests of the drive and the pitch using plastic balls in an indoor area. The test for the drive consisted of twenty trials with a brassie at a distance of twenty-five feet from a wall target. The target was eighteen feet wide and nine feet high. Scoring areas were three, five, seven, or ten points. The pitch test consisted of twenty trials with an eight iron from a hitting distance of twenty-eight feet from the target. The target area consisted of four concentric circles; the outer circle had a radius of seventy-five inches. Scoring areas were three, five, seven, or ten points. The reliability coefficients for sixty-four women classified as beginners were .75 for the drive test and .72 for the pitch test. Three judges' ratings of form correlated very low with obtained scores.

Cochrane (1960) constructed a battery of indoor tests to measure golfing ability. Plastic balls were used in the drive test, the mashie test, and the short approach test. A putting test using hard balls was also constructed. The reliability coefficient for the total test was .566 and validity, in relation to the handicaps of thirty-five subjects, was .604.

Chui (1965), in a study designed to determine the use of an electronic device as a teaching aid, developed two measures of golf skills that involved the seven iron and four iron. The testing area consisted of three concentric circles with a flag in the center. Diameters were twenty, sixty, and one hundred feet. Five strokes were allowed per subject. Three points were scored if the ball stopped in the inner circle; two points, the middle circle; and one point, the outer circle, provided the ball had first landed in any circle.

The quality of contact was also scored. A ball in the air beyond ten yards received one point; in a straight line, one additional point; and with a normal trajectory, another point. A total of six points was therefore possible on each trial. For the seven-iron item, the men were 85 yards from the flag and the women were 70 yards from the flag. The same target was used for the four-iron item. The men were 115 yards and the women were 100 yards from the flag. Reliability coefficients were .87 for the men and .84 for the women using the seven iron. Coefficients were .86 for the men and .75 for the women using the four iron.

Purdy and Stallard (1967) reported a five-iron drive test for distance and a ninety-yard test for accuracy in conjunction with a study concerning the acquisition of power and accuracy in the golf swing. The five-iron drive test consisted of hitting fifteen shots into a field lined at intervals of fifty feet. A tape measure was used to measure flight plus roll to the closest yard. For the ninety-yard test for accuracy, the subjects chose either a five, seven, or nine iron and hit fifteen shots at a flag surrounded by fifteen concentric circles. The radius was increased by fifteen feet for each circle. Trials were scored from one to fifteen points with fifteen representing the inner circle. The reliability coefficients for fifty-six college women were .82 for the five-iron test and .81 for the ninety-yard test. When correlated, a marked relationship existed between the two tests.

Brown (1969) developed a five-item battery of tests to evaluate golf skills. The chip test consisted of hitting fifteen shots into an athletic field that had three trapezoids marked on the ground. The hitter stood eighteen feet from the front of the center target. Trials were scored according to where the ball first hit; three, two, and one points were designated for the inner, middle, and outer trapezoids respectively.

The short pitch test consisted of hitting fifteen shots into an athletic field that had three concentric circles; radii were seven and one-half feet, fifteen feet, and twenty-two and one-half feet. The hitting line was sixty-five feet from the center of the target. Trials were scored according to where the ball first hit; three, two, and one points were designated for the inner, middle, and outer circles respectively.

The approach test consisted of hitting fifteen shots into an athletic field or golf course with dimensions and scoring the same as for the short pitch test except that yards were substituted for feet. Trials were scored according to where the ball came to rest.

The driving test consisted of hitting nine shots into an athletic field or golf course with two lines of markers placed 50 yards apart and located every 50 yards. Trials were scored according to where the ball came to rest. One point was scored for a shot short of the 100-yard line, two points if between the 100- and 150-yard lines, three points if between the 150- and 200-yard lines, four points if between the 200- and 250-yard lines,

and five points if beyond the 250-yard line. A ball stopping outside the 50-yard width markers was scored one less point than if in an adjacent zone.

The putting test consisted of putting twelve holes and the score was the total number of strokes taken. Six holes on a practice green were played twice, including two holes of about fifteen feet and four holes of about twenty feet. One hole was downhill, one was uphill, one broke to the left, one broke to the right, and two were level.

Reliability and validity coefficients are shown in Table I (Brown, 1969:4). The criterion for validation purposes was determined by adding the scores on nine holes of play.

TABLE I

RELIABILITY AND VALIDITY COEFFICIENTS FOR THE BROWN BATTERY

Tests	Reliability N	Reliability r	Validity N	Validity r
Chip Test	180	.74	86	.68
Short Pitch Test	148	.85	60	.76
Approach Test	155	.75	54	.65
Driving Test	104	.87	86	.73
Putting Test	58	.81	68	.71
Total Test Battery			134	.85

Gaskin and Porter (1972) designed self-testing outdoor events of the full swing, putt, and pitch-and-run shot. The full-swing event utilized a five iron. Twenty shots were taken into a testing area marked by flagsticks placed at a width of 30 yards and located every 20 yards. Trials were scored according to where the ball came to rest. The ball must have been airborne until past the 30-yard

line in order to score. One point was scored for a ball stopping between the 30- and 50-yard lines, two points if between the 50- and 70-yard lines, three points if between the 70- and 90-yard lines, and four points if between the 90- and 110-yard lines or beyond. A ball stopping outside the 30-yard width markers was scored one-half the points of the adjacent zone. Five players could hit at one testing area and players rotated with their scoring partners after ten shots.

The putting event consisted of nine holes; three holes were played from distances of six, twelve, and eighteen feet respectively. The pitch-and-run shot event consisted of hitting ten shots with a seven iron from a distance of no more than five yards from a regulation green. Balls were scored one, two, or four points respectively for landing on the green and rolling off, landing on the green and staying on, or stopping within a flagstick's length of the hole. Statistical data were not reported for the tests in this study.

Summary

One purpose of this review was to determine the relationship of the test items as compared to acceptable standards of test selection. The factors of validity, reliability, objectivity, and administrative feasibility are the usual criteria for test selection. Coefficients of validity and reliability were interpreted, with modifications, according to the levels summarized by Barrow and McGee (1971). Reliability coefficients of .80 or above and validity coefficients of .70 or above were considered acceptable in this study. Coefficients of objectivity are not usually reported in golf studies;

however, procedural explanations allow for an interpretation of the objectivity or subjectivity of scoring procedures. Specifically, tests were considered objective if scoring was done numerically, target areas were well defined, and special training for the scorers was not required. Administrative feasibility was interpreted in terms of whether or not a test could be administered to a group in a relatively short period of time. Specifically, tests were considered administratively feasible if individual measurements were not required after each trial and no more than thirty trials were designated.

Putting Tests

Reliability and/or validity coefficients were either not determined or not acceptable. The scoring procedures were objective but the tests would be time consuming to administer.

Approach Tests

The reliability and validity coefficients for the Nelson test were acceptable. The test was objective and administratively feasible. The reliability coefficient for the Smith test was acceptable; however, low validity was reported. The test was objective and administratively feasible. The West and Thorpe test was administratively feasible; however, reliability was low, subjective judgments were required, and validity measures were lacking.

Drive Tests

The reliability coefficients for the McKee test, using the two iron, were acceptable for the four items measured. Using the five iron, the reliability coefficients were acceptable for three of the four items measured. The test was objective; however, validity was not determined and the test would be time consuming to administer. The reliability coefficient for the Reese test was acceptable but validity was not determined. The test was objective and administratively feasible. The reliability coefficient for the distance score in the Benson test was acceptable and the validity coefficient for the distance and deviation scores in combination was acceptable. The test was administratively feasible; however, subjective judgments were required. The reliability coefficients for the Bevacqua test, using the five iron and three wood, were acceptable. Validity, however, was not determined. The test was objective and administratively feasible.

Test Batteries

Of the tests measured for reliability, coefficients were acceptable for the tests by Kelly, Vanderhoof, and Purdy and Stallard. Reliability coefficients were also acceptable for the outdoor distance driving test by Autrey; the seven-iron test for men and women and the four-iron test for men by Chui; and the short pitch test, driving test, and putting test by Brown.

Of the tests measured for validity, coefficients were acceptable for the two-iron distance test by Kelly; the drive test by Vanderhoof;

and the short pitch test, driving test, and putting test by Brown.

Of the tests measured for reliability and/or validity, the tests were objective except for the tests by Chui, which required subjective judgments.

Of the tests measured for reliability and/or validity that had coefficients that were acceptable, each was administratively feasible except for the outdoor distance driving test by Autrey, the two-iron distance test by Kelly, and the five-iron drive test by Purdy and Stallard.

CHAPTER III

PROCEDURES

The purpose of this study was to devise skill tests that indicate golf playing ability in group testing situations. An underlying assumption was that the fundamental elements of the game of golf may be classified as the drive, approach, and putt. Skill tests were devised for each element and analyzed according to selected basic criteria for test selection.

Selection of Subjects

All subjects were students enrolled at the University of North Carolina at Greensboro during the fall semester, 1972 or spring semester, 1973. The sample consisted of sixty subjects, forty-four enrolled in intermediate golf classes and sixteen volunteers with previous golf experience. A total of twenty-nine men and thirty-one women participated in the study.

The reason for selecting students enrolled in intermediate golf classes and requesting volunteers with previous golf experience was due to the selected method of establishing a criterion for validation purposes. The criterion consisted of game scores and a minimum of three rounds of play was arbitrarily established as a representative sample. Beginning golf classes were scheduled three times a week for sixty minutes each and intermediate classes

were scheduled twice a week for ninety minutes each. Game scores were to be obtained from play during regularly scheduled class periods. Beginning students would have difficulty in completing several rounds of play due to a lack of experience and the limited time period per class; therefore, they were not utilized in the study. This investigator believed that subjects with some golf experience would be more likely to provide stabilized scores on skill tests and game play than beginners. The study was limited to golfers with some previous experience in order to obtain data as reliable and valid as possible.

Development of Skill Tests

Distance and direction were the determining factors in assigning scoring values. Each test consisted of twenty trials, scoring was done by partners, regulation balls were used, and trials were scored according to where the ball came to rest. Complete test directions and sample scorecards are in the Appendix. A rationale for specific points and a clarification of ideas is presented in the following discussion.

Five-Iron Drive Test

Equipment. Although the drive is usually executed with a wood, a five iron was used in this test of the full swing. The greater distances that wood clubs usually produce could cause difficulty in scoring, require elaborate field markings, use more

space, and allow for fewer trials during the same testing period. The five iron is a versatile club designed for distance as well as accuracy and likely to be a familiar club to the intermediate golfer. The use of regulation balls and an outdoor grassy area for both the drive and approach tests was representative of an attempt to devise tests that were similar to the game situation.

Scoring. A study by Gaskin and Porter (1972) designated a lane width of 30 yards in a full-swing test using the five iron; balls landing outside the lane received half the score of the adjacent zone. Brown (1969) designated a lane width of 50 yards in a drive test using a wood; balls landing outside the lane received one point less than the adjacent zone. A lane width of 25 yards was designated for this test, with balls landing outside the lane scored as explained below. Twenty-yard increments seemed to provide adequate intervals for distance discrimination as evidenced by other studies using the five iron (Reese 1960, Bevacqua 1964, Gaskin and Porter 1972). The distances of 140 yards and 120 yards for men and women respectively, to mark the beginning of the highest scoring zones, were arbitrary choices.

The ball was scored where it came to rest and twenty trials were taken. Twenty trials were anticipated as being the maximum for the time allowed for administering each of the tests, yet adequate for establishing reliability. The same testing area was utilized for testing both men and women and was approximately 150 yards

long and 50 yards wide. Different scorecards were prepared to reflect a difference in point values for men and women. For men, a ball stopping between the 40- and 60-yard lines was scored one point; if between the 60- and 80-yard lines--two points; if between the 80- and 100-yard lines--three points; if between the 100- and 120-yard lines--four points; and if between the 120- and 140-yard lines--five points. A ball stopping beyond the 140-yard line was scored six points. A ball stopping outside the lane was scored two points less than the adjacent zone, but not less than zero. For women, a ball stopping between the 20- and 40-yard lines was scored one point; if between the 40- and 60-yard lines--two points; if between the 60- and 80-yard lines--three points; if between the 80- and 100-yard lines--four points; and if between the 100- and 120-yard lines--five points. A ball stopping beyond the 120-yard line was scored six points. A ball stopping outside the lane was scored two points less than the adjacent zone, but not less than zero. Trials were scored from 1 to 6 points and 120 points were possible for twenty trials.

Markings. Colored flags were used to mark the intersecting points of 20-yard increments at a lane width of 25 yards. White flags designated the hitting line; pink flags--20 yards; blue flags--40 yards; brown flags--60 yards; purple flags--80 yards; yellow flags--100 yards; green flags--120 yards; and red flags--140 yards. The field was marked in lime at the intersecting points prior to testing so the flags could be placed in a few minutes time. Connecting

lines between the markers might facilitate the scoring accuracy but the markers were found to be satisfactory to designate the scoring areas. The testing area is illustrated in Figure 1.

Approach Test

Equipment. A seven iron, eight iron, or nine iron was used in this test of the approach depending upon the subject's preference, experience, and availability of equipment. Each subject could use his own clubs or utilize clubs provided for class use which included the seven iron and nine iron. The use of the wedge is seldom a part of class instruction in a practical sense because the wedge is not usually provided in the allotted equipment for class use. Therefore, the wedge was not optional for this test.

Scoring. The lane width for this test was five yards with four-yard distance increments selected. A target flag was centered twenty yards from the hitting line in the highest scoring zone. The ball must have cleared the first ten yards in the air in order to score. A ball that cleared the first ten yards and stopped between the ten- and fourteen-yard lines was scored one point; if between the fourteen- and eighteen-yard lines--three points; if between the eighteen- and twenty-two-yard lines--five points; if between the twenty-two- and twenty-six-yard lines--three points; and if between the twenty-six- and thirty-yard lines--one point. A ball stopping outside the lane and going at least ten yards in the air was scored two points less than the adjacent zone, but not less than zero. Trials were scored from one to five points and one hundred points were possible for twenty trials.

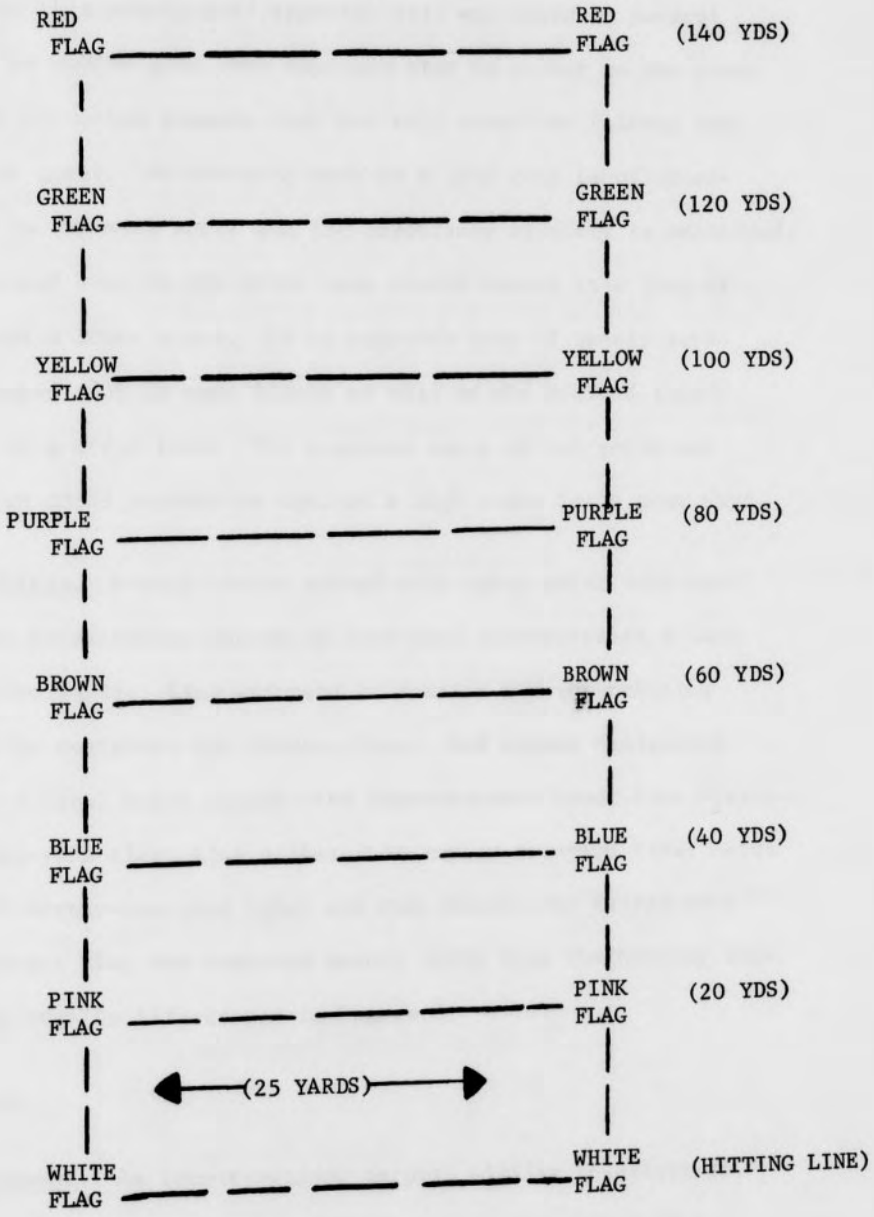


FIGURE 1

DIAGRAM OF THE TESTING AREA FOR THE

FIVE-IRON DRIVE TEST

The decision to require the ball to carry at least ten yards to score in this twenty-yard approach test was based on several factors. In course play, the approach shot is a shot to the green and proper execution demands that the ball clear the fairway and land on the green. An obstacle such as a sand trap is often encountered in approach shots and the importance of carry is magnified. A badly topped shot in the drive test should result in a loss of distance and a lower score. In an approach test of twenty yards, a badly topped shot is more likely to roll to the desired target area than in a drive test. The required carry of ten yards was therefore an added precaution against a high score for a poor shot.

Markings. Wooden stakes coated with spray paint were used to mark the intersecting points of four-yard increments at a lane width of five yards. Lime was used to connect the intersecting points and to designate the hitting line. Red stakes designated the ten-yard line; white stakes--the fourteen-yard line; blue stakes--the eighteen-yard line; blue stakes--the twenty-two-yard line; white stakes--the twenty-six-yard line; and red stakes--the thirty-yard line. A target flag was centered twenty yards from the hitting line. The testing area is illustrated in Figure 2.

Putting Test

Equipment. An indoor-outdoor carpet, similar to astroturf, was cut into two sections and two testing areas were prepared indoors. Although a putting cup is a good practice device, a putt

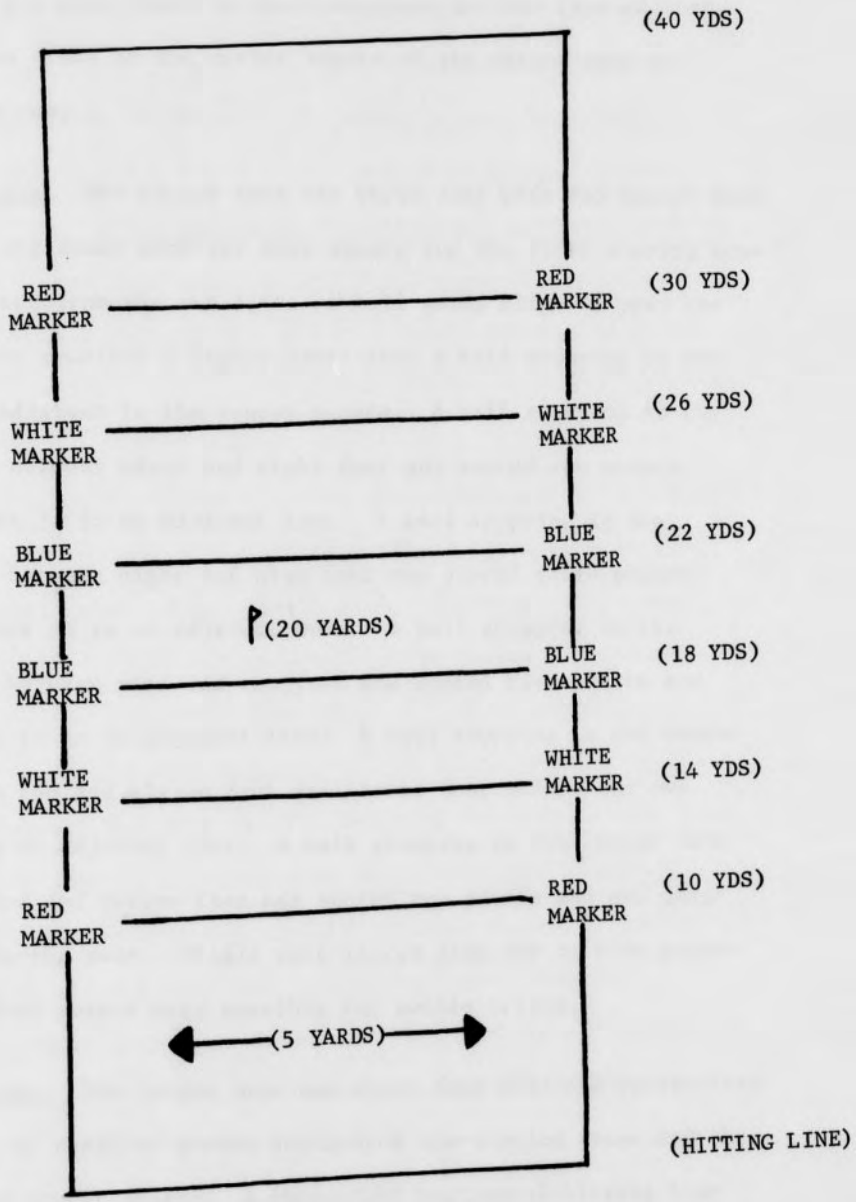


FIGURE 2

DIAGRAM OF THE TESTING AREA

FOR THE APPROACH TEST

that rims the cup could be thrown off line to the extent that the results of the putt would be misrepresented in this type of test. A circle was drawn in the center square of the target area to represent a cup.

Scoring. The target area was three feet wide and twelve feet long. Scoring zones were one foot square and the first scoring zone was seven feet from the end line. A ball going slightly past the center square received a higher score than a ball stopping in any other area adjacent to the center square. A ball stopping in the center lane between seven and eight feet was scored two points and one point if in an adjacent zone. A ball stopping in the center lane between eight and nine feet was scored three points and two points if in an adjacent zone. A ball stopping in the center lane between nine and ten feet was scored five points and three points if in an adjacent zone. A ball stopping in the center lane between ten and eleven feet was scored four points and two points if in an adjacent zone. A ball stopping in the center lane between eleven and twelve feet was scored two points and one point if in an adjacent zone. Trials were scored from one to five points and one hundred points were possible for twenty trials.

Markings. The target area was three feet wide and twelve feet long. Lines of chalk or powder designated the scoring areas and the circle in the center square. A three-foot horizontal hitting line was drawn six inches from the end line. A six-inch vertical dividing line was drawn eighteen inches from the side of the target area.

Balls were hit from within the six-inch area and each trial was taken from alternate sides of the vertical dividing line. Vertical lines five feet long were drawn at distances of one, two, and three feet from the side of the target area. Three-foot horizontal lines were drawn at distances of seven, eight, nine, ten, eleven, and twelve feet from the end line. A circle, four inches in diameter, was drawn in the center square nine feet from the hitting line. The testing area is illustrated in Figure 3.

Test Administration

Conferences were held with the instructors of intermediate golf classes during the fall semester, 1972 and spring semester, 1973. Cooperation was requested to allow one class period for the administration of the skill tests and to provide scores for a minimum of three rounds of nine holes of play for each subject.

Test directions and replicas of the target areas were given to the instructors the week prior to testing for distribution to each student. Testing was completed during the fourth week of October, 1972 for the two intermediate classes conducted during the fall semester, and completed on April 12, 1973 for the intermediate class conducted during the spring semester. Testing was conducted during a regularly scheduled class period of ninety minutes for each class.

Each testing session was conducted in a similar manner. Test directions were distributed and the subjects were randomly divided into three groups by distributing scorecards on

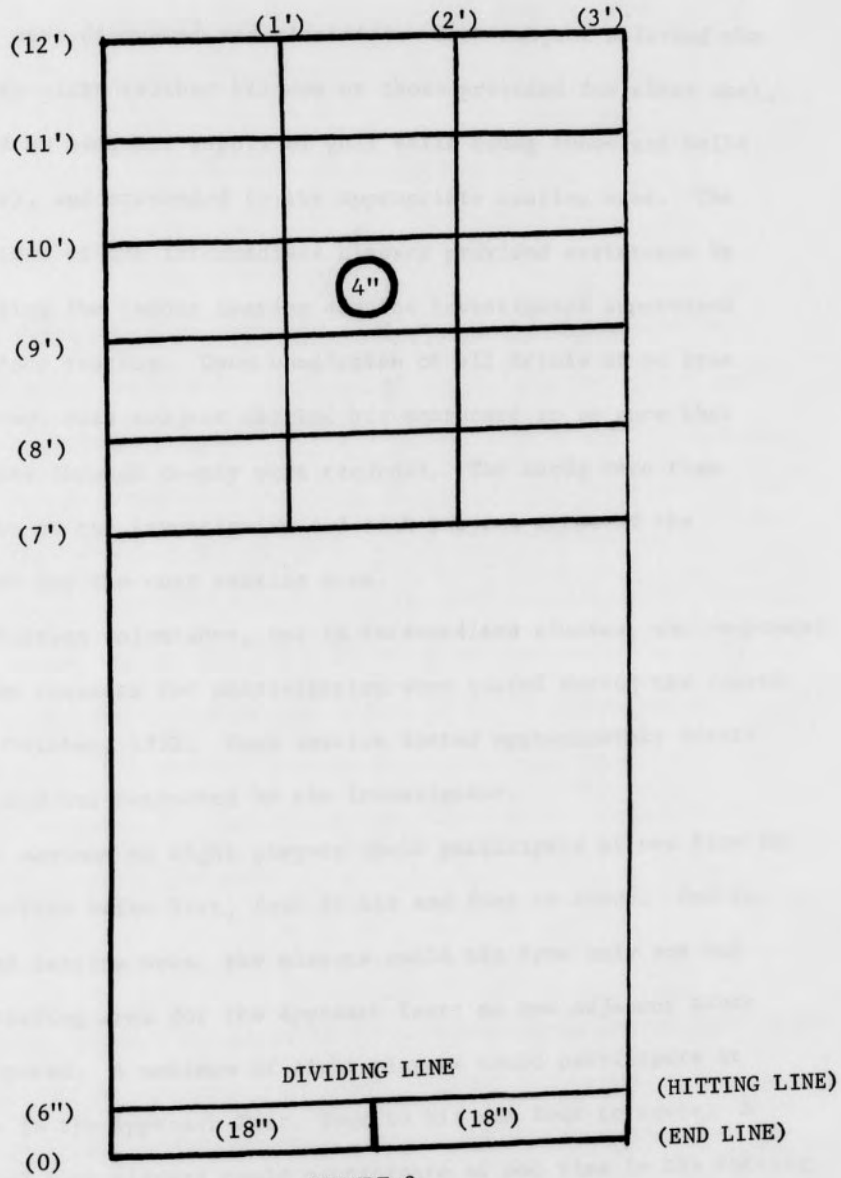


FIGURE 3

DIAGRAM OF THE TESTING AREA

FOR THE PUTTING TEST

five-by-eight-inch cards that contained replicas of either the drive, approach, or putting test areas. Test directions and scoring procedures were discussed and clarified. Each subject selected the necessary clubs (either his own or those provided for class use), gathered an adequate supply of golf balls (shag tubes and balls provided), and proceeded to the appropriate testing area. The instructors of the intermediate classes provided assistance by supervising the indoor testing and the investigator supervised the outdoor testing. Upon completion of all trials at an area by a group, each subject checked his scorecard to be sure that trials one through twenty were recorded. The cards were then turned in to the investigator and each subject received the scorecard for the next testing area.

Sixteen volunteers, not in intermediate classes, who responded to letter requests for participation were tested during the fourth week of October, 1972. Each session lasted approximately ninety minutes and was conducted by the investigator.

A maximum of eight players could participate at one time in the Five-Iron Drive Test, four to hit and four to score. Due to a limited testing area, the players could hit from only one end of the testing area for the Approach Test; so two adjacent areas were prepared. A maximum of eight players could participate at one time in the Approach Test, four to hit and four to score. A maximum of four players could participate at one time in the Putting Test on two prepared target areas, two to putt and two to score.

The testing areas were measured and lined prior to testing. Removable materials such as flags and stakes were inserted immediately prior to testing in approximately fifteen minutes time and removed upon completion of each testing session.

Collection of Data

Skill Test Scores

Scorecards were prepared that consisted of replicas of each of the testing areas with markings to indicate point values and yardage. Scoring was done by partners and each trial was recorded by placing the number of the trial in the appropriate space on the scorecard. Each subject checked his scorecard to be sure that trials one through twenty were recorded upon completion of testing at each area. Scorecards were collected at the end of each testing session. The final scores were computed by the investigator for each subject. The subject's score for each skill test was the sum of twenty trials. The sum of the odd-numbered and even-numbered trials was also determined.

Game Scores

Each subject participated in at least three and not more than twelve rounds of nine holes of play. Regulation scorecards for the University-owned par-thirty golf course were utilized for scoring purposes. Game scorecards were collected from the instructors of intermediate golf classes at the end of the fall or spring semester.

Game scorecards of volunteers, not in intermediate classes, were collected by the end of the fall semester, 1972. Each subject's score was computed by the investigator. An average score for nine holes of play was obtained by dividing the sum of strokes for all rounds completed by the number of rounds completed. The average number of rounds recorded was 3.94 for the women and 5.34 for the men.

Treatment of Data

A determination was made of the relationship of the skill test items as compared to acceptable standards of test selection in terms of objectivity, reliability, validity, and administrative feasibility. The standards stated below were also used in the review of literature to summarize previous skill tests.

Objectivity

A coefficient was not obtained; however, procedural explanations allowed for an interpretation of the objectivity or subjectivity of scoring procedures. Specifically, tests were considered objective if scoring was done numerically, target areas were well defined, and special training for the scorers was not required.

Administrative Feasibility

Administrative feasibility was interpreted in terms of whether or not a test could be administered to a group in a relatively short period of time. Specifically, tests were considered administratively

feasible if individual measurements were not required after each trial and no more than thirty trials were designated.

Reliability

To determine reliability for each test, the Split-Halves method using odd-numbered trials and even-numbered trials was utilized. Coefficients were obtained by the application of the Pearson Product-Moment method of correlation. The Spearman-Brown Prophecy formula was used to determine reliability coefficients for all twenty trials or to predict the coefficients for an additional number of trials. Coefficients of .80 or above were considered acceptable.

Validity

To determine validity for each test, game score averages were correlated with test scores. Coefficients were obtained by the application of the Pearson Product-Moment method of correlation. The Pearson Product-Moment method of correlation was also used to determine test intercorrelation coefficients. The Doolittle method for multiple correlations was used to determine validity coefficients for test combinations. Validity coefficients of .70 or above were considered acceptable.

CHAPTER IV

ANALYSIS OF DATA

The purpose of this study was to devise golf skill tests that indicate golf playing ability in group testing situations. An analysis was made of the objectivity, reliability, validity, and administrative feasibility of the test items.

A drive test, an approach test, and a putting test were devised and administered to sixty students enrolled at the University of North Carolina at Greensboro. Forty-four students were enrolled in intermediate golf classes and sixteen students were volunteers with previous golf experience. The tests were administered during the fall semester, 1972 and spring semester, 1973. A breakdown of subjects is in Table II.

TABLE II

NUMBER OF SUBJECTS PARTICIPATING IN THE STUDY

Subjects	Men	Women	Total
1. Intermediate Golf Class	3	12	15
2. Intermediate Golf Class	7	7	14
3. Intermediate Golf Class	13	2	15
4. Volunteers	6	10	16
	29	31	60

Objectivity

Coefficients were not obtained; however, an analysis was made in relation to previously stated criteria. Scoring was done numerically, target areas were well defined, and special training for the scorers was not required.

Administrative Feasibility

An analysis was made in relation to previously stated criteria. Individual measurements were not required after each trial and no more than thirty trials were designated. Twenty trials were taken on each test. The administration of all three tests to the classes tested required approximately ninety minutes, including the clarification of test directions.

Reliability

Reliability coefficients for each test were calculated for the sixty subjects. The Split-Halves coefficients and Spearman-Brown coefficients are listed in Table III.

TABLE III

COEFFICIENTS OF RELIABILITY FOR THE FIVE-IRON DRIVE TEST, APPROACH TEST, AND PUTTING TEST

Test Items	Trials Odd-Even	Stepped Up Twenty Trials
Five-Iron Drive Test	.93	.96
Approach Test	.79	.88
Putting Test	.48	.65

Odd-even reliability coefficients using ten trials were .93 for the Five-Iron Drive Test, .79 for the Approach Test, and .48 for the Putting Test. Coefficients using twenty trials were .96 for the Five-Iron Drive Test, .88 for the Approach Test, and .65 for the Putting Test. The coefficients using twenty trials were acceptable for the Five-Iron Drive Test and for the Approach Test. The Split-Halves coefficient of .93 suggests that ten trials might be sufficient for the Five-Iron Drive Test. The predicted number of trials necessary to reach an acceptable level of reliability (.81) on the Putting Test is forty-five. The reliability coefficients obtained on the Putting Test suggest that putting may be the most difficult element in the game of golf to repeat with consistency.

Validity

Validity coefficients for each test were calculated for the sixty subjects using game score averages as the criterion. Validity data are shown in Table IV.

TABLE IV

COEFFICIENTS OF VALIDITY USING GAME SCORE AVERAGES AS THE CRITERION

Test Items	Validity Coefficients
Five-Iron Drive Test	.71
Approach Test	.66
Putting Test	.59

Validity coefficients were .71 for the Five-Iron Drive Test, .66 for the Approach Test, and .59 for the Putting Test. The coefficient on the Five-Iron Drive Test was acceptable. The coefficient on the Approach Test was near the acceptable level; it would be questionable for validation purposes however. The validity coefficient on the Putting Test suggests that putting may be the least valid element in the evaluation of golfing ability.

Intercorrelation coefficients were calculated using the Pearson Product-Moment method of correlation. Intercorrelation data are listed in Table V.

TABLE V

COEFFICIENTS OF INTERCORRELATION FOR THE FIVE-IRON DRIVE TEST, APPROACH TEST, AND PUTTING TEST

Test Items	Intercorrelation Coefficients
Five-Iron Drive Test and Approach Test	.76
Approach Test and Putting Test	.61
Five-Iron Drive Test and Putting Test	.55

A high intercorrelation coefficient of .76 was obtained for the Five-Iron Drive Test and Approach Test in combination and the coefficients obtained for the other test combinations were modest (Weber and Lamb, 1970).

Multiple correlation coefficients were calculated using the Doolittle method and the data are shown in Table VI.

TABLE VI

COEFFICIENTS OF MULTIPLE CORRELATION FOR THE FIVE-IRON DRIVE TEST, APPROACH TEST, AND PUTTING TEST

Multiple Correlation Combinations*	Multiple Correlation Coefficients
R_0 123	.77
R_0 12	.73
R_0 13	.75
R_0 23	.70

*1. Five-Iron Drive Test
2. Approach Test
3. Putting Test

The multiple correlation coefficient of .77 for a three-item battery was acceptable; however, the predicted number of trials necessary to establish reliability on the Putting Test would make this item time consuming to administer. The Five-Iron Drive Test and Approach Test in combination, with a multiple correlation coefficient of .73, is a two-item battery with acceptable reliability coefficients for each item as tested.

Averages for game scores and test scores were calculated for the male and female subjects. Reliability and validity coefficients were also calculated and the data are presented in Table VII.

TABLE VII

DATA RELATING TO GAME SCORES AND TEST SCORES

FOR THE MALE AND FEMALE SUBJECTS

Items	Mean		Reliability***		Validity	
	Men*	Women**	Men	Women	Men	Women
Game Scores	36.45	47.72				
Five-Iron Drive Test	79.40	43.00	.96	.79	.73	.39
Approach Test	67.35	43.30	.89	.84	.66	.36
Putting Test	69.79	56.30	.75	.23	.59	.35

* N=29
**N=31
***20 Trials

As a group, those males participating in this study had attained a higher skill level than the females as evidenced by a comparison of game score averages. The males, as a group, also had higher average test scores on each respective test item than the females. Reliability and validity coefficients for the males were higher on each respective test item than for the females. The indication is that higher skilled subjects are more consistent in the performance of skill test items and reflect game play proficiency through skill testing more so than lower skilled subjects.

CHAPTER V

SUMMARY AND CONCLUSIONS

Summary

The purpose of this study was to devise golf skill tests that indicate golf playing ability in group testing situations. The skill tests were analyzed according to basic criteria for test selection.

Three skill test items were administered to sixty students enrolled at the University of North Carolina at Greensboro. Forty-four students were enrolled in intermediate golf classes and sixteen students were volunteers with previous golf experience. A total of twenty-nine males and thirty-one females participated in the study. The tests were administered during the fall semester, 1972 and spring, 1973.

A drive test, an approach test, and a putting test were devised to indicate golfing ability. Distance and direction were the determining factors in assigning scoring values. Each test item consisted of twenty trials, scoring was done by partners, regulation balls were used, and the ball was scored where it came to rest. The Five-Iron Drive Test consisted of hitting into a field that had scoring areas in 20-yard distance increments and a lane width of 25 yards. The highest scoring zone for the women was beyond the 120-yard line. For men, the highest scoring zone was beyond the 140-yard line.

The Approach Test consisted of hitting with either a seven, eight, or nine iron into a grassy target area twenty yards from the hitting line. Scoring areas were in four-foot increments and the lane width was five yards. A ball must have been airborne until past the ten-yard line in order to score. The Putting Test consisted of putting on a level carpet indoors. A simulated cup was nine feet from the hitting line and scoring areas were one foot square.

To determine objectivity, the skill test items were analyzed according to stated criteria related to scoring procedures. To determine administrative feasibility, the skill test items were analyzed according to stated criteria related to group testing.

To determine reliability, the Split-Halves method of using odd-numbered and even-numbered trials was utilized. Coefficients were obtained by the application of the Pearson Product-Moment method of correlation. The Spearman-Brown Prophecy formula was used to determine reliability coefficients for an entire test of twenty trials or to predict the coefficient for additional trials. Odd-even reliability coefficients were .93 for the Five-Iron Drive Test, .79 for the Approach Test, and .48 for the Putting Test. Reliability coefficients using twenty trials were .96 for the Five-Iron Drive Test, .88 for the Approach Test, and .65 for the Putting Test. Coefficients of .80 or above were considered acceptable.

To determine validity, game score averages were correlated with test scores. Validity coefficients were .71 for the Five-Iron Drive Test, .66 for the Approach Test, and .59 for the Putting

Test. The Pearson Product-Moment method of correlation was also used to determine test intercorrelation coefficients. The intercorrelation coefficient for the Five-Iron Drive Test and Approach Test was .76. The intercorrelation coefficient for the Approach Test and Putting Test was .61 and the coefficient for the Five-Iron Drive Test and Putting Test was .55. The Doolittle method for multiple correlations was used to determine validity coefficients for test combinations. The multiple correlation coefficient for a three-item battery was .77 and the coefficient for a two-item battery of the Five-Iron Drive Test and Approach Test was .73. The multiple correlation coefficient for a two-item battery of the Five-Iron Drive Test and Putting Test was .75 and the coefficient for the Approach Test and Putting Test was .70. Validity coefficients of .70 or above were considered acceptable.

Conclusions

The following conclusions were determined from the analysis of data:

1. Objectivity coefficients were not obtained; however, the Five-Iron Drive Test, Approach Test, and Putting Test were objective according to stated criteria related to scoring procedures.

2. The Five-Iron Drive Test, Approach Test, and Putting Test were administratively feasible according to stated criteria related to group testing and as demonstrated by the actual testing conducted as a part of this study.

3. The Five-Iron Drive Test was a reliable and valid measure of golfing ability. The Split-Halves reliability coefficient using ten trials was .93 and the coefficient using twenty trials was .96. The validity coefficient, in relation to the criterion of game score averages, was .71.

4. The Approach Test was a reliable measure of golfing ability. The reliability coefficient using twenty trials was .88. The validity coefficient of .66 would be questionable for validation purposes.

5. The high intercorrelation coefficient of .76 for the combination of the Five-Iron Drive Test and Approach Test indicates that one test item might be substituted for the other for testing purposes.

6. The reliability coefficient using twenty trials on the Putting Test was .65 and the validity coefficient was .59. The predicted number of trials necessary to reach an acceptable level of reliability (.81) was forty-five.

7. The multiple correlation coefficient of .77 for a three-item battery was acceptable; however, the predicted number of trials necessary to establish reliability on the Putting Test would make this item time consuming to administer.

8. The Five-Iron Drive Test and Approach Test in combination, with a multiple correlation coefficient of .73, was the only two-item battery with acceptable reliability coefficients for each item as tested. The coefficient of .73 for a battery of the Five-Iron Drive Test and Approach Test is only slightly higher than

the validity coefficient of .71 for the Five-Iron Drive Test as a single test item.

BIBLIOGRAPHY

BIBLIOGRAPHY

A. Books

Barrow, Harold M. and Rosemary McGee. A Practical Approach to Measurement in Physical Education. 2nd ed. Philadelphia: Lea & Febiger, 1971.

Lawther, John D. The Learning of Physical Skills. Englewood Cliffs, New Jersey: Prentice Hall, Inc., 1968.

Lema, Tony. Champagne Tony's Golf Tips. New York: McGraw-Hill Book Company, 1966.

Mathews, Donald K. Measurement in Physical Education. 3rd ed. Philadelphia: W. B. Saunders, Co., 1968.

Neilson, N. P. and C. R. Jensen. Measurement and Statistics in Physical Education. Belmont, California: Wadsworth Publishing Company, Inc., 1972.

Scott, M. Gladys and Ester French. Measurement and Evaluation in Physical Education. Dubuque, Iowa: Wm. C. Brown Co., 1959.

Weber, Jerome C. and David R. Lamb. Statistics and Research in Physical Education. Saint Louis: The C. V. Mosby Company, 1970.

Willgoose, Carl E. Evaluation in Health Education and Physical Education. New York: McGraw-Hill Book Company, Inc., 1961.

B. Periodicals

Bowen, Robert T. "Putting errors of beginning golfers using different points of aim," Research Quarterly, 39 (April, 1968), 31-35.

Brown, H. Stephen. "A Test Battery for Evaluating Golf Skills," Texas Association for Health, Physical Education, and Recreation Journal, (Spring Issue, 1969), 4-5, Courtesy of the National Golf Foundation.

Chui, Edward F. "A Study of Golf-O-Tron Utilization as a Teaching Aid in Relation to Improvement and Transfer," Research Quarterly, 36 (May, 1965), 147-152.

Clevett, Melvin A. "An Experiment in Teaching Methods of Golf," Research Quarterly, 2, N . 4 (December, 1931), 104-112.

Gaskin, Lynne and Nancy Porter. "Self-Testing-How Do You Measure Up?" DGWS Archery-Golf Guide-1972-1974, (1972), 137-140.

Kelliher, M. S. "Analysis of Two Styles of Putting," Research Quarterly, 34 (October, 1963), 344-349.

McKee, Mary E. "A Test for the Full Swinging Shot in Golf," Research Quarterly, 21 (March, 1950), 40-46.

Neale, D. C. and B. D. Anderson. "Accuracy of Aim with Conventional and Croquet-Style Golf Putters," Research Quarterly, 37 (March, 1966), 89-94.

Purdy, B. J. and M. L. Stallard. "Effect of two learning methods and two grips on the acquisition of power and accuracy in the golf swing of college women," Research Quarterly, 38 (October, 1967), 480-484.

West, C. and J. Thorpe. "Construction and validation of an eight-iron approach test," Research Quarterly, 39 (December, 1968), 1115-1120.

C. Unpublished Material

Autrey, Elizabeth P. "A Study of a Battery of Tests for Measuring Playing Ability in Golf." Unpublished Master's thesis, University of Wisconsin, 1937.

Benson, David W. "Measuring Golf Ability through Use of a Number Five Iron Test." Unpublished paper, Research Section, California AHPER Convention, April 7, 1963 as cited by:
Clarke, H. Harrison. Application of Measurement to Health and Physical Education. 4th ed. Englewood Cliffs, New Jersey: Prentice Hall, Inc. 1967, 317-318.

Bevacqua, Yvonne A. "A Comparison of the Influence of Two Club Progressions on the Ability of Beginning Golf Students to Hit a Golf Ball." Unpublished Master's thesis, University of Colorado, 1964.

Cochrane, June F. "The Construction of An Indoor Golf Skills Test As A Measure of Golfing Ability." Unpublished Master's thesis, University of Minnesota, 1960.

Coffey, Margaret. "Achievement Tests in Golf." Unpublished Master's thesis, State University of Iowa, 1946.

Gibb, Jesse L. "A Comparison of Putting Ability Following Instruction with Conventional and Concentrator Putters." Unpublished Master's thesis, Washington State University, 1964.

Kelly, Nancy. "A Study of a Battery of Tests for Measuring Golf Playing Ability of Beginners." Unpublished Master's thesis, University of Texas, 1944.

Lumpkin, Margaret. "Evaluation of a Putting Test in Golf." A Seminar Report, Wellesley College, 1945.

Nelson, Jack K. "An Achievement Test for Golf." Unpublished study, Louisiana State University, 1967, as cited by: Neilson, N. P. and C. R. Jensen. Measurement and Statistics in Physical Education. Belmont, California: Wadsworth Publishing Company, Inc., 1972, 238-239.

Olsen, Andrea C. "The Development of Objective Tests of the Ability of Freshmen and Sophomore Women to Drive and to Pitch a Plastic Golf Ball in a Limited Indoor Area." Unpublished Master's thesis, University of Colorado, 1958.

Randleman, Ronald J. "A Study Comparing the Effectiveness of Two Putting Styles for Beginning Golfers." Unpublished Master's thesis, University of Iowa, 1969.

Reese, Patsy A. "A Comparison of the Scores Made on an Outdoor and the Scores Made on an Indoor Golf Test by College Women." Unpublished Master's thesis, University of Colorado, 1960.

Smith, Barbara B. "The Effectiveness of Television Video Tape Instant Playback in Learning the Pitch and Run Shot in Golf." Unpublished Doctoral dissertation, University of North Carolina at Greensboro, 1968.

Vanderhoof, Ellen R. "Beginning Golf Achievement Tests." Unpublished Master's thesis, State University of Iowa, 1956.

Watts, Harriet. The Construction and Evaluation of a Target for Testing the Approach Shot in Golf." Unpublished Master's thesis, University of Wisconsin, 1942.

Williford, Edna Wolfe. "The Development of an Indoor Putting Test." Unpublished Master's thesis, University of North Carolina at Greensboro, 1970.

Wood, Janet I. "A Study for the Purpose of Setting Up the Specifications of a Golf Driving Cage Target and Test for the Mid-Iron and the Brassie Clubs." Unpublished Master's thesis, University of Wisconsin, 1933.

APPENDICES

APPENDIX A

LETTER REQUEST FOR PARTICIPATION

TEST DIRECTIONS

LETTER REQUEST FOR PARTICIPATION

Sophomore, junior, and senior physical education majors, as well as selected students with some golf experience, are being contacted through this letter to request your participation in a study concerning skill testing in golf.

Testing will be conducted the week of October 23rd (next week) in accordance with the time schedules listed below. You are requested to plan to attend one session which will consist of the administration of all three parts of the test. The attached sheet gives a description of the test items. Clubs and balls will be provided unless you desire to use your own clubs. In the event of inclement weather, indoor testing will be conducted.

Participants are desired who can furnish scores for at least 3 rounds of 9 holes of play on the UNC-G golf course. Rounds may be completed during the fall semester, 1972 and scores obtained from play since the beginning of the semester are acceptable. (Note: The junior majors have two recorded rounds already as part of class instruction).

Indicate when you plan to attend by placing a check (✓) in the appropriate space. If you anticipate a conflict in being able to participate in this study, please come at the time selected so other arrangements can be made as necessary.

Return the form to: SCOTTIE HUDSON

GRADUATE STUDENT MAILBOX
DOWNSTAIRS COLEMAN GYM

or

1557 WALKER AVE. 6
GREENSBORO, N.C. 27403

DETACH AND RETURN BY OCTOBER 23rd

NAME _____

CLASS _____

I can attend on the day and time indicated:

(Meet at putting green)

Oct. 24 Tuesday	Oct. 25 Wednesday	Oct. 26 Thursday	Oct. 27 Friday	Oct. 28 Saturday
2:00___	3:00 ___	2:00___	2:00___	10:00___
5:00___	5:00 ___	5:00___	5:00___	

TEST DIRECTIONS

Five-Iron Drive Test

1. Scoring is done by partners.
2. The test consists of twenty trials. Two practice hits may be taken prior to the first trial.
3. The ball is scored where it comes to rest (flight plus roll). A ball on a line receives the higher of the two scoring areas. Record the number of the trial in its corresponding scoring area on the scorecard.
4. A ball that is deflected by another ball does not count as a trial.
5. A swing and a miss counts as a trial.
6. Balls may be cleared from the target area when necessary and safe.
7. Do not total scores. When trials are completed, check to be sure that trials one through twenty are recorded.

Approach Test

1. Use a seven, eight, or nine iron.
2. Directions are the same as for the drive test with the following exception: Ball must clear the first ten yards to score. A ball that first touches the ground between the hitting line and the ten-yard line does not receive a score but is counted as a trial regardless of where the ball rolls and whether or not the ball is deflected by another ball.

Putting Test

1. Directions are the same as for the drive test.

Positioning

1. For the Five-Iron Drive Test and Approach Test, scoring partners each take ten trials from one of the numbered positions indicated on the scorecard by "FIRST 10." Partners move laterally for the second ten trials. The corresponding number indicated by "SECOND 10" represents the position for the second ten trials.
2. The scorers are behind the hitters on the Five-Iron Drive Test in a position that allows for proper scoring.
3. Players may be positioned at either or both ends of the testing area for the Approach Test.
4. For the Putting Test, each trial is taken from alternate sides of the dividing line. Scoring partners rotate duties after the first ten trials.

APPENDIX B

SCORECARD FOR MEN ON THE FIVE-IRON DRIVE TEST

SCORECARD FOR WOMEN ON THE FIVE-IRON DRIVE TEST

SCORECARD FOR THE APPROACH TEST

SCORECARD FOR THE PUTTING TEST

NAME _____
SCORE _____

FIVE-IRON DRIVE TEST
MEN

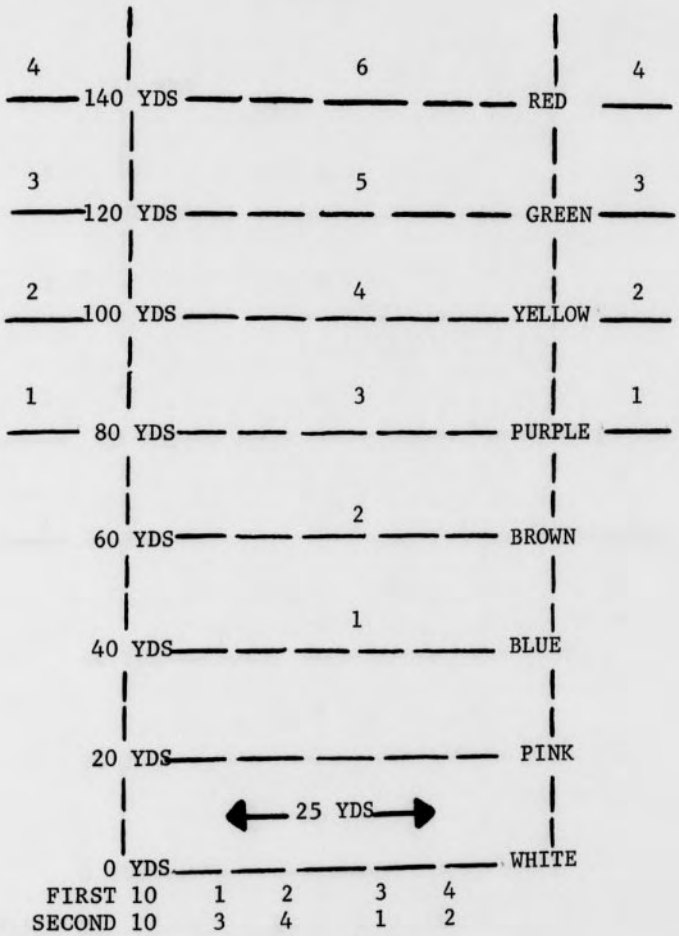


FIGURE 4

SCORECARD FOR MEN ON THE FIVE-IRON

DRIVE TEST

NAME _____
SCORE _____

FIVE-IRON DRIVE TEST
WOMEN

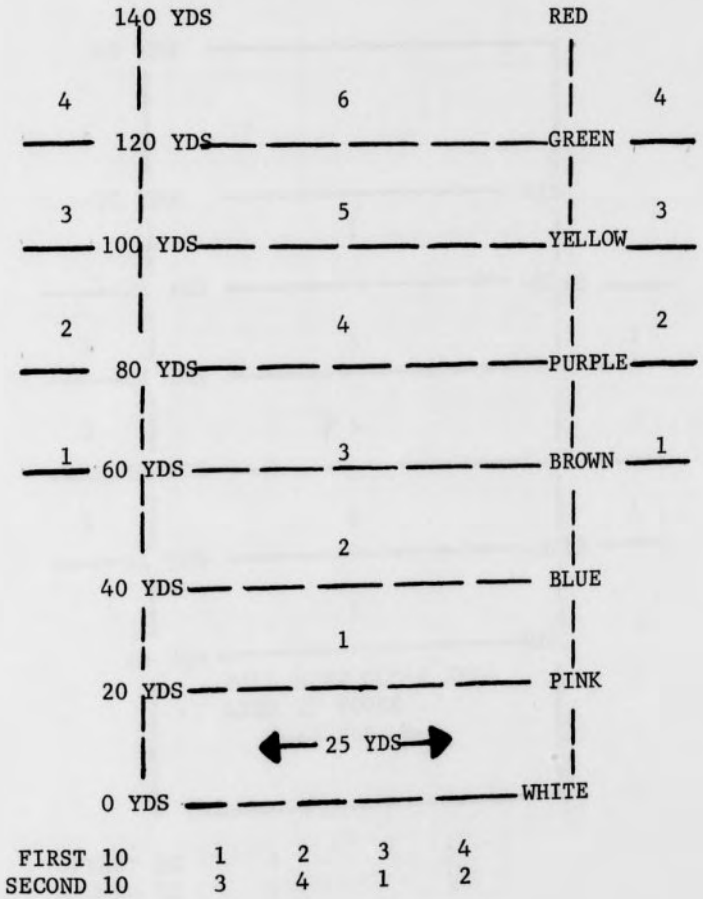


FIGURE 5

SCORECARD FOR WOMEN ON THE FIVE-IRON

DRIVE TEST

NAME _____
SCORE _____

APPROACH TEST

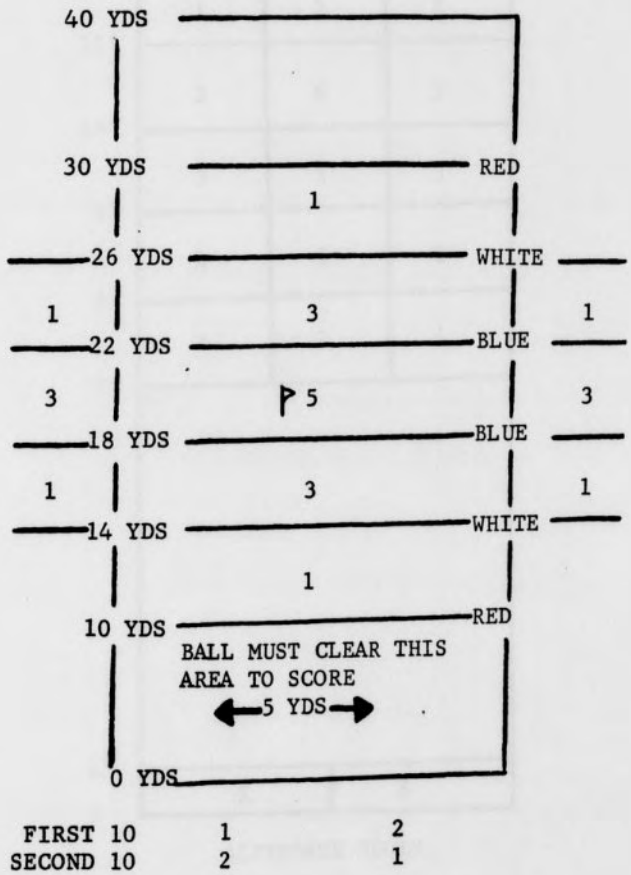
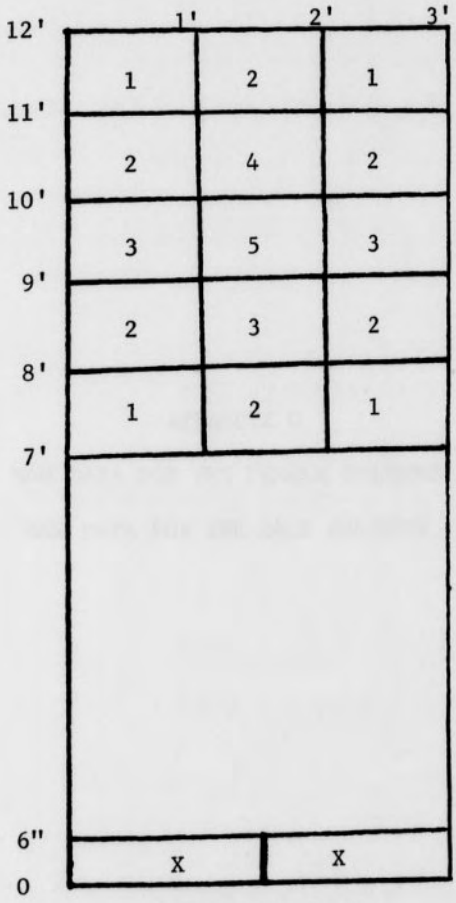


FIGURE 6

SCORECARD FOR THE APPROACH TEST

NAME _____
SCORE _____

PUTTING TEST



ALTERNATE SHOTS

FIGURE 7

SCORECARD FOR THE PUTTING TEST

APPENDIX C

RAW DATA FOR THE FEMALE SUBJECTS

RAW DATA FOR THE MALE SUBJECTS

TABLE VIII
RAW DATA FOR THE FEMALE SUBJECTS

Subjects	Test Scores								Game Scores
	Five-Iron Drive Test-Trials			Approach Test-Trials			Putting Test-Trials		
Females	Total	Odd	Even	Total	Odd	Even	Total	Odd	Even	Average
1.	57	29	28	52	28	24	60	31	29	40
2.	61	30	31	57	37	20	49	19	30	54
3.	41	24	17	43	16	27	60	35	25	35
4.	47	25	22	49	24	25	70	32	38	37
5.	52	25	27	23	4	19	47	22	25	46
6.	62	31	31	40	21	19	59	26	33	41
7.	44	19	25	54	27	27	75	37	38	46
8.	54	24	30	43	18	25	64	30	34	45
9.	47	18	29	18	5	13	52	28	24	55
10.	42	23	19	57	33	24	71	30	41	45
11.	37	18	19	26	17	9	46	16	30	57
12.	36	24	12	64	27	37	66	28	38	41
13.	48	24	24	46	18	28	51	23	28	54
14.	21	10	11	36	17	19	64	34	30	57
15.	40	21	19	29	10	19	53	32	21	61
16.	17	13	4	26	12	14	58	27	31	52
17.	45	24	21	44	23	21	40	21	19	38
18.	48	25	23	23	10	13	54	21	33	49
19.	45	20	25	25	14	11	55	26	29	46
20.	17	9	8	10	5	5	56	32	24	55
21.	33	19	14	33	15	18	56	29	27	54
22.	51	27	24	61	29	32	67	33	34	35
23.	58	25	33	58	27	31	52	28	24	51
24.	43	19	24	70	38	32	61	34	27	46
25.	38	13	25	66	32	34	70	37	33	55
26.	39	16	23	47	20	27	57	29	28	43
27.	31	14	17	48	29	19	51	22	29	52
28.	47	20	27	60	34	26	50	27	23	47
29.	42	22	20	49	16	33	50	34	16	55
30.	31	18	13	11	5	6	61	29	32	49
31.	56	29	27	76	36	40	70	33	37	42

TABLE IX
RAW DATA FOR THE MALE SUBJECTS

Subjects	Test Scores									Game Scores
	Five-Iron Drive Test-Trials			Approach Test-Trials			Putting Test-Trials			
Males	Total	Odd-Even		Total	Odd-Even		Total	Odd-Even		Average
1.	81	38	43	52	29	23	72	37	35	38
2.	63	28	35	67	33	34	75	38	37	35
3.	58	29	29	63	35	28	56	26	30	37
4.	94	48	46	82	42	40	83	41	42	29
5.	58	26	32	61	27	34	75	39	36	34
6.	104	50	54	69	35	34	65	34	31	32
7.	61	28	33	43	25	18	69	33	36	33
8.	42	21	21	49	26	23	64	35	29	41
9.	42	25	17	41	30	11	61	33	28	44
10.	116	56	60	83	40	43	87	42	45	31
11.	107	52	55	80	38	42	77	41	36	36
12.	111	59	52	78	38	40	73	37	36	37
13.	95	48	47	82	40	42	74	38	36	37
14.	74	35	39	72	36	36	82	43	39	38
15.	108	53	55	75	33	42	72	40	32	35
16.	56	26	30	46	27	19	71	39	32	40
17.	21	11	10	30	16	14	78	43	35	46
18.	65	36	29	50	21	29	57	28	29	37
19.	42	20	22	48	23	25	53	31	22	44
20.	79	41	38	88	44	44	77	42	35	38
21.	75	36	39	86	44	42	66	34	32	39
22.	82	40	42	82	40	42	69	40	29	36
23.	77	38	39	52	24	28	46	22	24	38
24.	92	46	46	80	38	42	72	37	35	31
25.	24	8	16	52	28	24	51	28	23	40
26.	85	37	48	78	44	34	69	31	38	36
27.	100	52	48	94	44	50	72	34	38	36
28.	103	51	52	76	37	39	80	37	43	32
29.	98	51	47	90	46	44	73	34	39	32